Cross-Layer Modeling Framework for Energy-Efficient Resilience

***Pradip Bose*[*], *David Brooks*[+], *Subhasish Mitra*[!], *Karthick Rajamani*[**], *Mircea Stan*[#], *Kevin Skadron*[##], *Gu-Yeon Wei*[+]**

*IBM T. J. Watson Research Center, Yorktown Heights, NY
**IBM Austin Research Laboratory, Austin, TX
[+]Dept. of Electrical Engineering and Computer Science, Harvard University, Cambridge, MA
[!]Dept. of Electrical Engineering and Dept. of Computer Science, Stanford University, Stanford, CA
[##]Dept. of Computer Science; [#]Dept. of Electrical and Comp. Engg., University of Virginia, Charlottesville, VA

Abstract: *We describe a novel cross-layer, resilience-focused integrated modeling framework. This is targeted to help define ultra energy-efficient embedded systems in the post-14nm CMOS design era, without compromising system-level resilience. The targeted application domain is represented by the suite of applications and kernels announced as part of the ongoing PERFECT program sponsored by DARPA MTO.*

Keywords: cross-layer modeling; embedded systems; resilience optimization; energy efficiency.

Introduction

The system-level efficiency (i.e. GFLOPS/watt) targets for the DoE-sponsored Exascale program and the DoD (DARPA MTO) sponsored PERFECT program are quite similar. In either case, the general architectural paradigm being pursued by the R&D community at the chip-level is a many-core design, possibly heterogeneous in terms of the compute elements, with the supply voltage pushed down as low as possible. Aspects of 3D packaging technology, coupled with concepts like near-memory computing to help meet the performance and efficiency targets are being pursued by various teams. Across all proposals, the issue of system-level resilience is a critical one that often gets ignored in concept-phase definitions of power-aware chip- and system-level (micro)architectural proposals. In this paper, we describe our ongoing effort (under the DARPA PERFECT program) to develop a cross-layer, resilience-focused integrated modeling framework. The goal is to demonstrate feasibility of ultra-high energy efficiency, while maintaining present-day system resilience levels in the post-14nm CMOS technology regime. Our initial thrust is on developing an analytical modeling framework that enables the study of fundamental power-performance-reliability trade-offs, while serving as a validation reference for more detailed, cycle-accurate modeling infrastructure in future phases of the PERFECT program.

Cross-Layer Modeling Strategy

Figure 1 depicts the integrated, cross-layer system modeling concept as pursued in the IBM-led project titled: "Efficient Resilience in Embedded Computing."

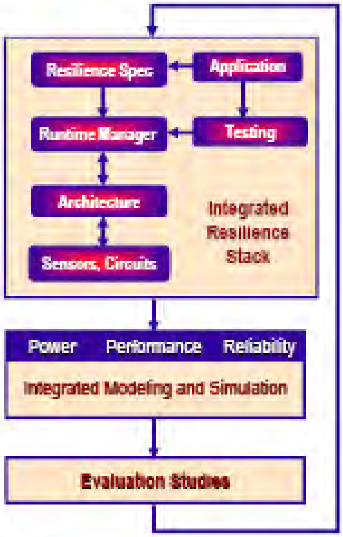


Figure 1. Cross-Layer Modeling Concept

The system is modeled as a layered stack ranging from lowest level instantiations of hardware sensors, circuits and packaging at the processor chip level up through system-level architectural constructs (including memory), system software and the user-level application.

The modeling strategy is orchestrated through seven tasks as outlined below:

T1: The top-level cross-layer resilience optimization framework, with associated modeling environment.
T2: R-API, a resilience-aware application programming interface for adaptive resilience provisioning. (R-API is the smart user interface through which the application developer interacts with the PEARL modeling framework, as described later).
T3: Efficient memory resilience, taking advantage of low-leakage storage-class technologies as appropriate.
T4: Ultra-efficient microarchitecture to provide low power resilience solution support at the system level.
T5: Optimal voltage point selection – static and dynamic.
T6: Resilient circuits, technology and packaging.
T7: Resilient resource management (for robust power control; energy-secure operation).

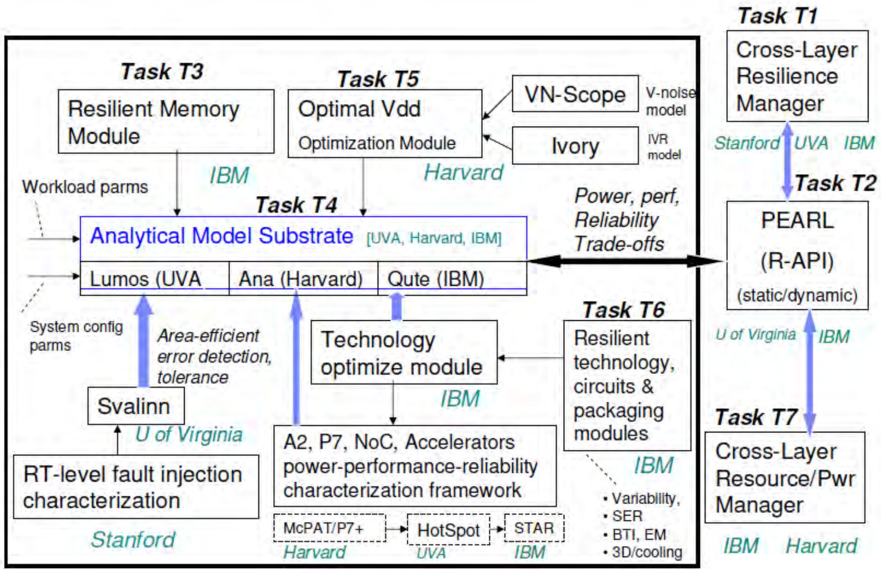


Figure 2. Overall Modeling Framework

Overall Modeling Framework

Figure 2 depicts the overall modeling framework under development in this project. The software is built around the substrate analytical power-performance model for a multi-core, heterogeneous microprocessor chip as pursued under Task T4. This model currently has three distinct modules: (a) the *Lumos* model developed at the University of Virginia [1]; (b) the *Ana* model developed at Harvard University [2]; and, (c) the *Qute* model developed at IBM Research [3]. The first two are both developed around basic analytical formalisms based on Amdahl's Law. Qute is an analytical model based on queuing theory, in which task arrivals and service times are modeled using user-selectable probability distribution functions. This is useful in modeling computational environments in which hundreds or thousands of accelerator threads are spawned off in support of main computational functions and are executed on a massively multi-threaded engine (e.g. a GPGPU sub-system).

The power model is driven by a technology optimization module (built around prior work by David Frank et al. [11] which captures the effect of technology scaling in arriving at optimized multi-core configurations in the post-14nm CMOS regime. Two different general purpose cores are currently modeled: (a) the POWER7 (P7) class core [4], known for its high single-thread performance as well as high throughput efficiency for general purpose codes; and (b) the A2 core [5], with quad SIMD floating point unit that provides very high GFLOPS/watt efficiency in IBM's Blue Gene/Q (BGQ) supercomputer.

In addition, because *sort* and *fft* are key kernels of interest within the PERFECT application suite, area-efficient accelerators geared towards those specific kernels are also targeted for inclusion in our modeling library. Feeding into the substrate processor power-performance model are models for memory (task T3), voltage sensitivity (task T5) and abstractions of circuit-technology-package level parameters (task T6). As part of task T5, two novel point tools called *VN-Scope* [6] and *Ivory* [7] have been developed. The former addresses the problem of modeling voltage noise by first developing a power delivery network (PDN) model. The Ivory tool models the effect of integrated voltage regulators (IVRs) – which is of increasing importance in future processor designs.

The *Svalinn* model [8] was developed to capture the area-reliability trade-offs inherent in the choice of specific error detection and tolerance mechanisms in microprocessor design. The model is being integrated with learning derived from RT-level statistical fault injection experiments conducted in a standalone research project [9]. This latter set of experiments yields accurate characterization of the effects of bit-flips at the latch or flip-flop level as they percolate up the system stack. This experimental characterization also helps formulate (or validate) the statistical fault injection model adopted for application-level "derating" calculations that are used in the overall soft error rate (SER) estimations derived as part of task T6, and as required to drive the PEARL/R-API framework (see the next section).

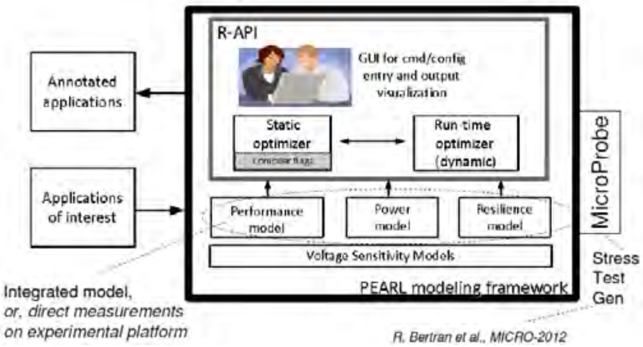
Task T5 is actually closely tied to task T6, in which circuit-level abstractions of low-level failure mechanisms are pursued as part of the overall modeling hierarchy. The effects of package- and cooling parameters on overall system efficiency (GFLOPS/watt) are also captured through the T6 modeling activity.

There is a separate, ongoing effort to develop a calibrated power model for the baseline POWER7 chip using the open-source *McPAT* [10] toolset; this will feed a POWER7-specific temperature model built around *HotSpot* [12]. Application-driven chip thermal maps will feed into adapted versions of IBM's lifetime reliability model called *STAR* [13], which is being retargeted to the POWER7 chip design. The specific failure models within STAR's lifetime reliability modeling capability include: electromigration (EM), negative bias temperature instability (NBTI) and time-dependent dielectric breakdown (TDDB). These aging (or wearout) models are also linked to the voltage sensitivity effects of task T5.

The two cross-layer optimization tasks (T1 and T7) involve the development of static and dynamic optimizers that collectively maximize delivered performance per watt, while meeting stipulated system resilience targets. The PEARL/R-API framework discussed in the next section utilizes T1 and T7 algorithms, while also taking advantage of the analytical power-performance-reliability reliability trade-off capability of models derived from tasks T2 through T6. Overall, the modeling framework depicted in Figure 2 reflects a close-knit collaborative effort between IBM and its three university partners (Stanford, Harvard and University of Virginia) in the PERFECT project.

PEARL/R-API: Application Development Facility

In this section, we focus on one of our key innovations in the overall cross-layer modeling exercise. This is the PEARL/R-API framework that we briefly alluded to in the previous section. Figure 3 shows the high-level functional block diagram of the software architecture of PEARL, which stands for: *Power Efficient and Resilient Embedded Processing with Real-Time Constraints*. The user is able to utilize a smart graphical user interface (R-API) to: (a) select and compile an application from the repository; and (b) characterize it offline to understand the phase-wise behavior in terms of energy usage, performance and vulnerability to errors.

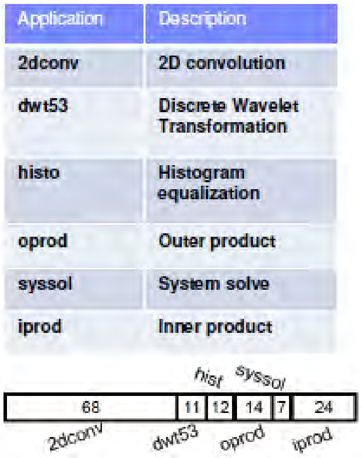


R. Bertran et al., MICRO-2012

Figure 3. PEARL Framework

The dynamic aspect of PEARL then allows the user to deploy the instrumented application on the simulated embedded processor system in a manner that allows the run-time manager to adjust power control knobs (e.g. dynamic voltage-frequency scaling, DVFS). The goal of the run-time manager is to minimize power consumption, while maintaining system resilience targets (on average) and meeting real-time performance targets.

The integrated performance, power and resilience models are nothing but the analytical modeling toolkit described in the previous section. However, in lieu of such models, the PEARL framework can also be used by invoking direct measurement tools that are part of the actual hardware platform that the PEARL software is executing on. The *MicroProbe* [14] facility allows the user to generate focused stress test suites that can be used to test the limits of the static and dynamic optimization.



Application	Description
2dconv	2D convolution
dwt53	Discrete Wavelet Transformation
histo	Histogram equalization
oprod	Outer product
syssol	System solve
iprod	Inner product

hist
syssol
68
11
12
14
7
24
2dconv
dwt53
oprod
iprod

Figure 4. Workflow Consisting of Six PERFECT Applications

Application level resilience characterization profiles are generated using the application-level fault injection (AFI) tool that is part of our SER estimation tool chain. The sensitivity of errors (transient or hard) with respect to operational voltage is captured through the "voltage sensitivity" module (see Figure 2) which is derived through empirical analysis performed using IBM's internal pre-silicon processor reliability analysis framework.

Figure 4 shows a workflow constructed from six key applications that are part of the recently announced PERFECT application suite. The nominal execution times (in seconds) on a 4.1 GHz POWER7+ processor system are indicated in the workflow schematic below the table in Figure 4.

Figure 5 shows static (average) SER vulnerability profiles across the range of individual applications in the above workflow. These are obtained using fault injection experiments performed using our AFI tool (see prior section). The injections can result in one of four effects: *masked* completely (no effect), software *crash*, *hung* or no forward progress and *silent data corruption* (SDC).

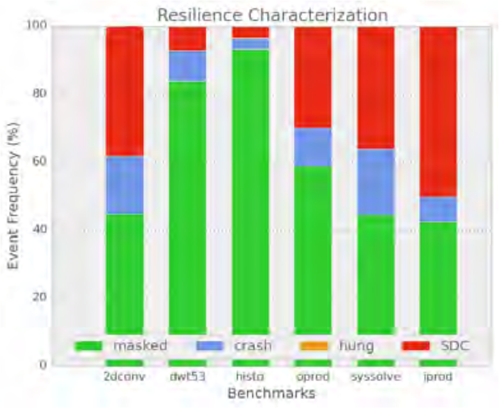


Figure 5. Static Resilience Characterization Using AFI

The static optimizer within PEARL (see Figure 3) can be invoked to assign optimal settings of voltage-frequency points to each application segment within the workflow. The objective function is performance per watt, under stipulated constraints of maximum power and SER resilience. The former constraint serves as a guard against over-clocking beyond a certain limit; and the latter constraint serves as a guard against setting the voltage below a certain limit – since SER increases with reduction in voltage. The dynamic optimizer within PEARL is able to handle in-field uncertainties caused by the harsh operating conditions in which mobile, airborne embedded systems (represented by unmanned aerial vehicles) have to operate.

Acknowledgments

This work is sponsored by Defense Advanced Research Projects Agency, Microsystems Technology Office (MTO), under contract no. HR0011-13-C-0022. The views expressed are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government. This document is: Approved for Public Release, Distribution Unlimited.

References

1. L. Wang, K. Skadron, "Implications of the power wall: dim cores and reconfigurable logic," *IEEE Micro*, vol. 35, issue 3, pp. 40-49, Sept/Oct 2013.

2. C-N. Tseng, D. Brooks, "Analytical latency-throughput model for future power-constrained multi-core processors," *Proc. Workshop on Energy-Efficient Design (WEED),* held in conjunction with International Symposium on Computer Architecture, June 2012.

3. N. Madan, A. Buyuktosunoglu, P. Bose, M. Annavaram, "A case for guarded power gating for multi-core processors," Proc. *Int'l. Symp. on High Performance Computer Architecture (HPCA),* Feb. 2011, pp. 291-300.

4. B. Sinharoy et al. "IBM POWER7 multicore server processor". *IBM Journal of Research and Development*, vol. 55, no. 3, May 2011.

5. R. A. Haring, M. Ohmacht, T. W. Fox, M. K. Gschwind et al., "The IBM Blue Gene/Q compute chip," *IEEE Micro*, vol. 32, issue 2, March/April 2012.

6. X. Zhang, T. Tong, S. Kanev, S. Lee, G-Y. Wei, D. Brooks, "Characterizing and Evaluating Voltage Noise in Multi-Core Near-Threshold Processors," *Proc. International Symposium on Low Power Electronics and Design (ISLPED).* September 2013.

7. Ivory – a new toolset from Harvard University that models power delivery subsystems using IVRs; publication pending.

8. L. Szafaryn, B. H. Meyer, and K. Skadron. "Evaluating the overheads of soft error protection mechanisms in the context of multi-bit errors at the scope of a processor core," *IEEE Micro* special issue on Reliability Aware Design, 33(4):56-65, July-Aug. 2013.

9. H. Cho, S. Mirkhani, C-Y. Cher, J. A. Abraham, S. Mitra, "Quantitative evaluation of soft error injection techniques for robust system design," *Proc. ACM/IEEE Design Automation Conference (DAC),* June 2013.

10**.** S. Li, J. H. Ahn, R. D. Strong, J. B. Brockman, D. M. Tullsen, N. P. Jouppi, "McPAT: An integrated power, area and timing modeling framework for multicore and manycore architectures," *Proc. 42nd Int'l. Symp. on Microarchitecture* (MICRO-42), December 2009.

11. L. Chang and D. J. Frank, "Technology Optimization for High Energy-Efficiency Computation," *IEDM 2012* Short Course.

12. K. Skadron, K. Sankaranarayanan, S. Velusamy, D. Tarjan, M.R. Stan, and W. Huang. "Temperature-Aware Microarchitecture: Modeling and Implementation." *ACM Transactions on Architecture and Code Optimization,* 1(1):94-125, Mar. 2004.

13. J. Shin, V. Zyuban, Z. Hu, J. Rivers, P. Bose, "A Framework for Architecture-Level Lifetime Reliability Modeling," *Proc. Int'l. Conf. on Dependable Systems and Networks (DSN),* 2007, pp. 534-543.

14. R. Bertran, A. Buyuktosunoglu, M. Gupta, M. González, P. Bose: Systematic Energy Characterization of CMP/SMT Processor Systems via Automated Micro-Benchmarks, *Proc. Int'l. Symp. on Microarchitecture, MICRO, 2012*, pp. 199-211.